

Bahasa yang digunakan adalah Bahasa Daerah Kota Cirebon yang dibuat dalam bentuk excel yang nantinya akan diinputkan pada algoritma IBM metode I. Beberapa contoh kalimat yang dibangun adalah sebagai berikut:

	Bahasa Cirebon	Arti
0	Isun dolanan geni	saya bermain api
1	Ira nggoleti wadon bae	kamu mencari perempuan terus
2	Watu ning kali abot pisan	batu di sungai berat sekali
3	Kebriben carae ngeliwet sega ?	bagaimana caranya memasak nasi ?
4	umah iku duwur pisan	rumah itu tinggi sekali

Gambar 1 Bahasa Cirebon beserta artinya dalam Bahasa Indonesia

Percobaan telah dilakukan menggunakan suatu kata Bahasa Cirebon dengan arti sebenarnya pada Bahasa Indonesia, hasil yang diperoleh adalah sebagai berikut:

```
1 print(round(com_ibm1.translation_table["api"]["geni"], 2) )
0.47

1 print(round(com_ibm1.translation_table["saya"]["isun"], 2) )
0.98

1 print(round(com_ibm1.translation_table["tidur"]["turu"], 2) )
0.74
```

Gambar 2 Hasil percobaan pada 3 kata beserta arti sebenarnya

Hasil dari percobaan ketiga kata menggunakan arti sebenarnya pada sebenarnya, memiliki probabilitas yang berbeda. Perbedaan ini dikarenakan sedikit banyaknya suatu kata tersebut dengan artinya pada suatu kalimat corpus. Kata “geni” yang memiliki arti “api” pada corpus yang telah dibuat hanya memiliki 1 data sehingga hasilnya relatif kecil sedangkan kata “isun” yang berarti “saya” memiliki nilai yang hampir sempurna hal ini dikarenakan pada corpus yang dibangun kata “isun” yang memiliki arti “saya” terdapat pada 6 data.

<pre>1 counterCrb = 0 2 counterInd = 0 3 for i in range(len(corpus)): 4 crb = str(corpus.iloc[i,0]).lower().split() 5 ind = str(corpus.iloc[i,1]).lower().split() 6 counterCrb += len([w for w in crb if w == "geni"]) 7 counterInd += len([w for w in ind if w == "api"]) 8 counterCrb, counterInd</pre>	<pre>1 counterCrb = 0 2 counterInd = 0 3 for i in range(len(corpus)): 4 crb = str(corpus.iloc[i,0]).lower().split() 5 ind = str(corpus.iloc[i,1]).lower().split() 6 counterCrb += len([w for w in crb if w == "isun"]) 7 counterInd += len([w for w in ind if w == "saya"]) 8 counterCrb, counterInd</pre>
(1, 1)	(6, 6)

Gambar 3 Perbandingan banyaknya kata dalam corpus

Hal ini dikarenakan IBM Model 1 hanya melakukan translasi terhadap lexical dan urutan katanya. Sehingga semakin banyak data suatu kata dengan arti yang dimaksud nilai probabilitas juga akan semakin besar, dan kebalikannya. Pernyataan ini sudah terbukti oleh percobaan yang telah dilakukan pada gambar ke-2